UC San Diego

UC San Diego Previously Published Works

Title

Exploring the relationship between polygenic risk for cannabis use, peer cannabis use and the longitudinal course of cannabis involvement.

Permalink

https://escholarship.org/uc/item/5rj3h1ms

Journal

Addiction (Abingdon, England), 114(4)

ISSN

0965-2140

Authors

Johnson, Emma C
Tillman, Rebecca
Aliev, Fazil
et al.

Publication Date

2019-04-01

DOI

10.1111/add.14512

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

ADDICTION
SSA SOCIETY FOR THE STUDY OF ADDICTION
RESEARCH REPORT
doi:10.1111/add.14512

Exploring the relationship between polygenic risk for cannabis use, peer cannabis use and the longitudinal course of cannabis involvement

Emma C. Johnson[1], Rebecca Tillman[1], Fazil Aliev[2,3], Jacquelyn L. Meyers[4], Jessica E. Salvatore[2], Andrey P. Anokhin[1], Danielle M. Dick[2,5], Howard J. Edenberg[6], John R. Kramer[7], Samuel Kuperman[7], Vivia V. McCutcheon[1], John I. Nurnberger Jr[8], Bernice Porjesz[4], Marc A. Schuckit[9], Jay Tischfield[10], Kathleen K. Bucholz[1] & Arpana Agrawal[1]

ABSTRACT

Background and aims Few studies have explored how polygenic propensity to cannabis use unfolds across development, and no studies have yet examined this question in the context of environmental contributions such as peer cannabis use. Outlining the factors that contribute to progression from cannabis initiation to problem use over time may ultimately provide insights into mechanisms for targeted interventions. We sought to examine the relationships between polygenic liability for cannabis use, cannabis use trajectories from ages 12–30 years and perceived peer cannabis use at ages 12–17 years. **Design** Mixed-effect logistic and linear regressions were used to examine associations between polygenic risk scores, cannabis use trajectory membership and perceived peer cannabis use. **Setting** United States. **Participants** From the Collaborative Study on the Genetics of Alcoholism (COGA) study, a cohort of 1167 individuals aged 12–26 years at their baseline (i.e. first) interview. **Measurements** Key measurements included life-time cannabis use (yes/no), frequency of past 12-month cannabis use, maximum life-time frequency of cannabis use, cannabis use disorder (using DSM-5 criteria) and perceived peer cannabis use. Polygenic risk scores (PRS) were created using summary statistics from a large (n = 162 082) genome-wide association study (GWAS) of cannabis use. **Findings** Three trajectories reflecting no/low (n = 844), moderate (n = 137) and high (n = 186) use were identified. PRS were significantly associated with trajectory membership [P = 0.002–0.006, maximum conditional R^2 = 1.4%, odds ratios (ORs) = 1.40–1.49]. Individuals who reported that most/all of their best friends used cannabis had significantly higher PRS than those who reported that none of their friends were users [OR = 1.35, 95% confidence interval (CI) = 1.04, 1.75, P = 0.023]. Perceived peer use itself explained up to 11.3% of the variance in trajectory class membership (OR = 1.50–4.65). When peer cannabis use and the cannabis use PRS were entered into the model simultaneously, both the PRS and peer use continued to be significantly associated with class membership ($P < 0.01$). **Conclusions** Genetic propensity to cannabis use derived from heterogeneous samples appears to correlate with longitudinal increases in cannabis use frequency in young adults.

Keywords Cannabis use, externalizing behaviors, high-risk sample, peer influence, polygenic risk score, trajectories.

Correspondence to: Emma C. Johnson, 314-273-1873, Department of Psychiatry, Washington University School of Medicine, 660 South Euclid, CB 8134, St Louis, MO 63110, USA. E-mail: emma.c.johnson@wustl.edu
Submitted 22 June 2018; initial review completed 13 August 2018; final version accepted 16 November 2018
Please see list of authors' affiliations at the end of the paper.

INTRODUCTION

The growing controversy regarding cannabis legalization in the United States [1] is based in part on the question of whether increased access is associated with escalations of both use and misuse [2], with the latter currently affecting approximately 6% of the population [3]. Longitudinal studies have classified young cannabis users into those who remain casual users, those who transition to moderate levels of use and remain stable, those who show initial increases followed by declines in use and, importantly, those who demonstrate accelerated use and progression to problem use [4–11]. Outlining the factors that contribute to the likelihood of progression to problem use might provide insights into targets for intervention.

Cannabis use and misuse are heritable (h^2 = 50–70% of the variation). Several genome-wide association studies (GWAS) have attempted to identify loci that might contribute to this heritable variation [12–18]. For cannabis use, the largest published study to date (n = 184 765

© 2018 Society for the Study of Addiction

individuals of European descent [19]; results used here on $n = 162\,082$; see Supporting information for details) identified four independent genome-wide significant loci and found a genome-wide single nucleotide polymorphism (SNP) heritability of 10%, suggesting that the aggregated effects of common SNPs captured a sizeable proportion of the heritability of cannabis use. Polygenic risk scores (PRS) offer a complementary approach to the study of such aggregated effects [20]. In brief, a PRS is a person-specific index of genetic propensity to a trait (e.g. cannabis use); PRS are constructed by multiplying the effect size from a discovery GWAS by the number of risk alleles that an individual possesses at that SNP. PRS approaches are widely used in psychiatric genetics, including substance use and dependence, and can be used to assess whether genetic risk for one disorder or trait is associated with aspects of the same trait or with a correlated disorder/trait [21,22]. For instance, one study found that PRS for schizophrenia risk predicted cannabis use in individuals with bipolar disorder [23]. However, few studies have explored how genetic propensity to cannabis initiation (i.e. cannabis PRS) influences patterns of cannabis use across development.

In addition to genetic risk, affiliations with cannabis-using peers are believed to be among the leading contributors to persistent cannabis use [8,11,24,25]. However, results from longitudinal samples remain mixed (e.g. [7]). While peer use is readily viewed as an 'environmental' agent of risk, it can also represent heritable aspects of underlying behavior, with at least one study suggesting a heritability of 25–28% for general peer group deviance, a broad measure including peer marijuana use [26]. That study also found that approximately 50–78% of the genetic variance in peer group deviance was attributable to genetic factors related to cannabis use [27–29]. Another study [30] reported that the heritability of perceived peer alcohol use ranged from 7% at age 12–14 to 38% by age 18, and that the relationship between peer alcohol use and one's own alcohol use was attributable to genetic factors with a correlation of 0.83. Taken together, these observations raise the possibility that polygenic risk for cannabis use may interface with peer cannabis use in several possible ways, ranging from a main effect to a potential interactive effect. To our knowledge, these hypotheses remain untested.

To understand more clearly the role of genetic propensity and peer use in the longitudinal course of cannabis use, we used data on 1167 individuals of European descent who were part of a large longitudinal study of the genetics of addictions. We first identified trajectories of cannabis use frequency, and then examined whether trajectory class membership was related to (a) cannabis use PRS and/or (b) perceived peer cannabis use when the subject was aged 12–17 years. We also examined whether the relationship between polygenic risk, perceived peer use and trajectory membership could be explained by an interaction model where perceived peer use moderated the influence of polygenic risk on trajectory membership. Results from these analyses can provide a framework for how genetic liability and peer use might interface to shape the developmental unfolding of cannabis use.

METHODS

Participants

The Collaborative Study on the Genetics of Alcoholism (COGA) study recruited alcohol-dependent probands through substance use treatment programs at seven sites throughout the United States. Probands and their family members were invited to participate, resulting in an over-representation of densely affected multiplex pedigrees. Control families (two parents and three or more offspring over the age of 14) were also selected from a variety of community sources (e.g. driver license registries). The institutional review boards for all seven data collection sites, and additional data analysis sites, approved the study [31].

For the current analyses, data from a cohort of 3618 individuals ('September 2017' data freeze) who were aged 12–26 years at their baseline (i.e. first) interview and comprised the longitudinal component of COGA, were used [32]. Briefly, participants were offspring of COGA families, with 61.6% having one parent with alcohol use disorder. Since 2004, participants have been interviewed every 2 years with the same structured interview; follow-up interviews are ongoing. We included only subjects with GWAS data and of European–American (EA; as verified by genotype) descent to match the ethnicity of the discovery GWAS [13] ($n = 1897$); of these individuals, 1840 had non-missing data for relevant variables. For the longitudinal growth curve analyses, a further reduction in sample size resulted from subsetting on those who were EA, had GWAS data and had three or more assessments, including the baseline assessment (final analytical $n = 1167$). When compared to the larger subset of 1840 individuals, those with three or more assessments did not vary on any demographic or cannabis-related characteristics, suggesting that selection for those with greater than or equal to three assessments did not significantly bias findings (Supporting information, Table S1).

Assessment

All individuals were interviewed using a version of the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA [33,34]) with individuals aged < 18 years administered a child version (C-SSAGA [33,34]).

- Life-time cannabis use was coded using all assessment responses to an item on whether they had ever used cannabis (response: yes or no).

© 2018 Society for the Study of Addiction

- Frequency of past 12-month cannabis use was also recorded as the response at each interview to an item querying how often the participant had used cannabis in the past 12 months; the range was from 0 to 935. Data were winsorized to remove outliers (> 3 standard deviations) at each age and were binned into 31 categories in 20-unit intervals. The interval length was chosen to capture variation in the data and allow for model optimization (sensitivity analyses with a 10-unit interval were also conducted; see Supporting information).
- Maximum life-time frequency of cannabis use was the maximum reported frequency of past 12-month cannabis use; this variable was log-transformed before analysis due to being right-skewed.
- Cannabis use disorder (CUD) was coded using DSM-5 criteria [35] (without the requirement for clustering of symptoms).
- Perceived peer cannabis use was coded as the response to an item: 'When you were 12–17, how many of your best friends used marijuana?' (0 = none, 1 = a few, 2 = most, 3 = all; categories representing 'most' and 'all' were combined, as the latter was only endorsed by 35 individuals); to minimize recall bias, peer use reported at the last assessment was used for those aged 12–17 (n = 818), and at the assessment closest to age 12–17 for subjects aged 18 and older (349, although 91% were aged 18–21 years).

Genotypic data

Members of COGA's prospective cohort were genotyped as a part of multiple initiatives on different Illumina and Affymetrix arrays. The reported pedigree structure was assessed using a pruned set of 1 519 440 SNPs. In total, 6 881 872 SNPs passed quality control and data cleaning thresholds and were available for analysis (see details in Supporting information).

Polygenic risk for cannabis use

Effect sizes and effect alleles were derived from genome-wide summary statistics from a large GWAS meta-analysis of 162 082 individuals, all of European ancestry (characteristics of discovery GWAS [19] in Supporting information). PRS were created for each COGA individual of genetically verified European descent with SNPs meeting increasingly lenient P-value thresholds from the discovery GWAS (from $P_T < 0.0001$ to $P_T < 0.50$). Details are provided in Supporting information but briefly, for each COGA individual, effect sizes from the discovery GWAS by Pasman *et al.* were multiplied by the number of effect alleles for each SNP, and then averaged across all SNPs within a certain P-value threshold (e.g. $P_T < 0.10$) (e.g. tuning parameter [36]) in the discovery GWAS to create one score per individual for that P_T. This P_T threshold is not reflective of significance of the PRS in a traditional statistical sense (i.e. $P < 0.05$). Instead, it is predicated on the assumption of a high degree of polygenicity, which has been found to be true for most complex traits [37]; therefore, SNPs that do not reach stringent genome-wide significance cut-offs (typically $P <$ 5e–8) in the discovery GWAS are still predicted to make small but incremental and additive contributions to risk liability for the outcome [20,38,39].

Covariates

Sex, age at first (i.e. baseline) and last interview were included as covariates. Two additional covariates were also included. First, three principal components, reflecting continuous variation in genetic ancestry were derived from all the GWAS data (details in Supporting information) and included to account for subtle ancestral differences [40]. Secondly, the type of genotyping array used for each individual was included as a covariate in order to control for potential differences in genomic content, quality control or imputation (see Supporting information, Table S1 for descriptive data).

Statistical analyses

Estimation of trajectories

Only subjects with GWAS data and cannabis frequency of use data available at three or more assessment waves (n = 1167) were included in the growth mixture models. Latent class growth analysis (LCGA) with a zero-inflated Poisson model in MPlus version 8 [41] was used to assign these individuals to classes that were derived using cannabis past-year frequency of use categories from each of the up to seven interviews (baseline to 12-year follow-up). Age at assessment was used as the analytical unit (i.e. the x-axis). Analysis details are available in Supporting information.

PRS analyses

The PRS were standardized (using the 'scale' function in R) before analysis. Mixed-effect logistic regression models were used to test the association between PRS (at varying P_T) and trajectory class membership (using pairwise comparisons between classes), the associations between peer use and trajectory class membership, between the PRS and peer use and also to test whether an interaction between PRS and peer cannabis use predicted trajectory membership, while accounting for all second-order interaction terms (see [42]). All analyses included the family identifier and recruitment site as random effects (family nested within site).

All the above analyses were conducted in R [43]. To assess model fit and the relative amount of variance explained by the PRS, we used the 'MuMIn' package in R

© 2018 Society for the Study of Addiction

to calculate both marginal and conditional R^2 for each mixed model [43]. We use the conditional R^2 to select the most predictive P_T (see Supporting information), but report both statistics for the most predictive PRS threshold. The proportion of variance attributable to the PRS (Δconditional R [2]) was estimated as the difference between the conditional R^2 for a model with covariates alone and the model that included covariates and the PRS [i.e. conditional R^2 (full model) – conditional R^2 (model without PRS)]. The use of ΔR [2] (typically Nagelkerke's pseudo-R^2 for binary traits [20]) as an index of the most predictive PRS relates to its role as an index of predictor efficacy [38,39,44], such that the addition of the PRS to a model improves the fit of the model, thus indicating unique variance attributable to the PRS, over and above covariates. As peer use was restricted to recall at age 12–17, we did not test whether trajectory membership influenced future peer use. The Bonferroni-corrected significance threshold for the PRS analyses was set at 0.0019 (corrected for 27 tests: three class comparisons × nine PRS thresholds), while the significance threshold was set at $\alpha < 0.05$ for the remaining analyses. In addition, to overcome concerns that uncertainty in class membership might have influenced results, we re-ran analyses for the most predictive PRS threshold using the BCH approach in MPlus [45]. In this approach, the LCGA model is fitted to data and weights are assigned to likelihood of membership in each class while simultaneously examining between-class differences in PRS and accounting for the effect of covariates on class membership (see Supporting information).

Role of externalizing behaviors

To examine whether cannabis use PRS represented a general propensity to externalizing behaviors, we examined their association with (a) the thrill/adventure-seeking and the disinhibition subscales from Zuckerman's sensation scale (from baseline assessments; for adults [46]; Russo's modified sensation-seeking scale for children [47]; standardized) and (b) with a life-time diagnosis of conduct disorder from the SSAGA.

Negative control analyses

As a negative control, we also tested whether the PRS significantly predicted height at baseline, a trait not expected to be genetically associated with cannabis use.

RESULTS

Trajectories of recent cannabis use

As shown in Supporting information, Fig. S1, three classes were identified as the three-class model had a lower Bayesian Information Criterion (BIC) than the two-class solution, the Lo–Mendel–Rubin adjusted likelihood ratio test (LMR-ALRT) P-value for the four-class solution ($P = 0.1002$) was not significant, and the entropy for the three-class solution (0.917) was high (fit statistics in Supporting information, Table S2; parameter estimates for the best-fitting model in Supporting information, Table S3). Classification probabilities were high (0.975, 0.928, 0.977). Sensitivity analyses with 10-unit intervals of cannabis use frequency were similar (Supporting information, Table S2). Broadly speaking (Table 1), the classes represented (a) users who consistently used cannabis infrequently during the entire period of follow-up, and included never users of cannabis, that we termed the 'no-low' use class ($n = 844$); (b) another class that included individuals with very high frequency of initial use that continued to escalate during the follow-up period and remained elevated at the final assessment, that we termed the 'high' use class ($n = 186$); and (c) a class that included escalating use that involved similar high use at baseline but a less steep increase in use during the follow-up, that we termed the 'moderate' use class ($n = 137$). Also, as shown in Table 1, those in the high and moderate use trajectories were significantly more likely to be male, have used cannabis at an earlier age and meet criteria for a life-time history of cannabis use disorder (CUD) as well as conduct disorder.

Associations between cannabis use PRS and overall cannabis use in the sample

In the analytical sample ($n = 1167$) we found no evidence that the cannabis PRS was associated with a binary measure of life-time cannabis use ($P = 0.111$), nor with frequency of use at baseline (in ever-users, $P = 0.390$), frequency of use at last assessment (in ever-users, $P = 0.513$) or maximum frequency of cannabis use (in ever-users, $P = 0.090$). However, the cannabis use PRS was associated with life-time history of DSM-5 CUD ($P = 0.028$) but was no longer significant in the subset of ever-users ($P = 0.090$, Supporting information, Table S4). The pattern of association with cannabis use was similar when individuals with fewer than three assessments ($n = 1840$) were studied, although in this larger sample the PRS was associated with maximum frequency of cannabis use ($P = 0.013$) and with DSM-5 cannabis use disorder in both the full sample ($P = 0.005$) and in ever-users ($n = 1144$, $P = 0.014$).

Cannabis use PRS predicting cannabis use trajectories

The cannabis use PRS was significantly associated with cannabis use trajectory class membership (Table 2). At the most significantly associated PRS threshold of $P_T < 0.1$,

© 2018 Society for the Study of Addiction

Table 1 Characteristics of European–American individuals in classes representing high, moderate and no-low cannabis use frequency.

	High-use class (n = 186)		*Moderate-use class (n = 137)*		*No-low-use class (n = 844)*		*High versus no-low*	*High versus moderate*	*Moderate versus no-low*
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*OR (95% CI)*	*OR (95% CI)*	*OR (95% CI)*
Baseline age	15.74	3.00	16.14	3.29	15.43	3.12	1.01 (0.96, 1.07)	0.96 (0.89, 1.03)	1.07 (1.00, 1.13)*
Age at last assessment	23.71	4.07	25.04	4.05	23.91	4.38	0.97 (0.93, 1.01)	0.92 (0.87, 0.98)**	1.06 (1.01, 1.11)*
Age at first cannabis use	15.08	2.08	15.98	1.99	17.87	2.76	0.61 (0.56, 0.68)**	0.80 (0.71, 0.90)**	0.71 (0.708, 0.715)**
Frequency of use at baseline	87.66	153.63	32.12	89.62	0.72	4.15	1.07 (1.04, 1.10)**	1.004 (1.002, 1.007)**	1.11 (1.07, 1.15)**
Frequency of use at last assessment	278.28	193.02	120.57	139.19	5.21	25.58	1.08 (1.06, 1.10)**	1.007 (1.005, 1.009)**	1.033 (1.026, 1.041)**
	%	*n*	*%*	*n*	*%*	*n*	*OR (95% CI)*	*OR (95% CI)*	*OR (95% CI)*
Male gender	67.20	125	64.96	89	41.94	354	3.00 (2.99, 3.00)**	1.10 (0.68, 1.77)	2.90 (1.90, 4.43)**
Life-time cannabis use	100.00	186	100.00	137	51.66	436	–	–	–
Cannabis use disorder	84.40	157	69.34	95	7.10	60	127.14 (41.19, 392.41)**	2.45 (1.39, 4.31)**	39.73 (19.40, 81.37)**
Conduct disorder diagnosis	34.41	64	21.17	29	7.23	61	7.38 (4.61, 11.80)**	1.96 (1.17, 3.28)*	4.09 (2.28, 7.31)**

*$P < 0.05$; **$P < 0.01$. SD = standard deviation; CI = confidence interval; OR = odds ratio.

the cannabis use PRS explained approximately 1.4% of the conditional variance in high versus no-low class membership (2.30% of the marginal variance); for every unit increase in PRS, membership in the high versus no-low class increased by an odds ratio (OR) of 1.40 [95% confidence interval (CI) = 1.13, 1.74] (Supporting information, Fig. S2; full results for all thresholds in Table 2, all covariates in Table 3). Cannabis PRS also explained 3.6% of the conditional variance in high versus moderate class membership, although this comparison did not survive Bonferroni correction (OR = 1.49, 95% CI = 1.12, 1.97, $P = 0.006$). There was no evidence that cannabis use PRS was associated with height at baseline ($P = 0.730$). Results from the BCH approach identified identically significant differences in mean PRS across the high class when compared with the moderate and the no-low class (Supporting information, Table S5).

Peer cannabis use predicting cannabis use trajectories

Of the 1162 individuals with peer use data available, 57.5, 28.3 and 14.2% reported that none, few and most or all of their close peers used cannabis. Perceived peer cannabis use explained up to 11.3% of the variance in trajectory class membership (ORs = 1.50–4.65). When peer cannabis use and the cannabis use PRS were entered into the model simultaneously, the association between the cannabis use PRS and membership in high versus moderate class was only slightly attenuated (OR = 1.46, 95% CI = 1.09, 1.94; $P = 0.010$), as was the association with the high versus no-low class comparison (OR = 1.34, 95% CI = 1.07, 1.68, $P = 0.012$). Peer cannabis use was independently and significantly associated with all three comparisons in these models that also included the PRS as a predictor (both high versus no-low and moderate versus no-low $P < 0.001$, high versus moderate $P = 0.017$).

Cannabis use PRS predicting peer cannabis use

Those who reported that most or all of their best friends used cannabis had significantly higher PRS than those who reported that 'none' of their best friends used cannabis (most significant OR = 1.38, 95% CI = 1.07, 1.78, $P = 0.012$). Other comparisons (e.g. none versus few: $P = 0.799$; few versus most or all, $P = 0.096$) did not significantly differ from each other on cannabis PRS (details in Supporting information).

Role of externalizing behaviors

The disinhibition scale score was significantly associated with both trajectory membership and all peer cannabis

© 2018 Society for the Study of Addiction

Table 2 Results from mixed-effect logistic regression models predicting cannabis use trajectory class membership by polygenic risk scores for cannabis initiation. All models controlled for age at baseline, age at last assessment, sex, the first three ancestry principal components and array type, and site and family id were included as nested random effects. The *n* SNPs column is the minimum number of SNPs included in each PRS threshold (some individuals had fewer SNPs included in the score due to missing genotypes). The PRS *P*-value threshold of $P_T < 0.1$ is shown in bold type, as this PRS was the most significant threshold associated with belonging to the high use class compared to both the no-low use class and the moderate use class, and this PRS threshold explained the most conditional variance (see Fig. S2). Thus, this PRS was used in all analyses reported in the main paper.

		Moderate class versus no-low class			*High class versus no-low class*			*High class versus moderate class*		
P-value threshold (P_T)	*n SNPs*	*Beta*	*SE*	*P*	*Beta*	*SE*	*P*	*Beta*	*SE*	*P*
*P*5	298 678	0.054	0.128	0.676	0.242	0.106	0.022	0.185	0.132	0.162
*P*4	258 733	0.042	0.121	0.727	0.284	0.107	0.008	0.222	0.137	0.106
*P*3	211 705	0.034	0.128	0.792	0.306	0.111	0.006	0.256	0.136	0.060
*P*2	157 496	0.022	0.124	0.857	0.271	0.106	0.011	0.239	0.138	0.083
***P*1**[a]	**92 504**	**−0.055**	**0.121**	**0.651**	**0.339**	**0.109**	**0.0018**	**0.396**	**0.145**	**0.006**
*P*05	52 656	−0.116	0.116	0.315	0.263	0.104	0.011	0.390	0.140	0.005
*P*01	14 102	−0.122	0.115	0.289	0.011	0.097	0.910	0.109	0.132	0.408
*P*001	2166	0.066	0.111	0.553	−0.003	0.097	0.978	−0.081	0.123	0.508
*P*0001	372	0.131	0.111	0.240	0.076	0.099	0.445	−0.099	0.132	0.451

[a]Corresponding results for PRS $P_T < 0.1$ using the BCH approach are in Supporting information, Table S5. PRS = polygenic risk scores; SNP = single nucleotide polymorphism; SE = standard error.

Table 3 Associations between cannabis initiation polygenic scores and cannabis use trajectories. The most significant PRS is reported, which was defined with a *P*-value threshold of $P_T < 0.1$ (see Table 2 for results for all thresholds).

	Moderate versus no-low			*High versus no-low*			*High versus moderate*		
	Beta	*SE*	*P*	*Beta*	*SE*	*P*	*Beta*	*SE*	*P*
Cannabis use PRS	−0.055	0.121	0.651	**0.339**	**0.109**	**0.0018**	0.396	0.145	0.006
Sex	**1.117**	**0.222**	**5.03e–07**	**1.037**	**0.194**	**9.79e–08**	0.015	0.258	0.955
Age at baseline	−0.046	0.065	0.482	0.140	0.057	0.014	0.207	0.078	0.008
Age at last assessment	0.101	0.050	0.042	−0.116	0.043	0.007	**−0.225**	**0.064**	**4.12e–04**
Principal component 1	−276.537	124.649	0.027	−224.026	129.429	0.083	−51.340	219.127	0.815
Principal component 2	−81.749	82.385	0.321	101.019	101.375	0.319	69.935	168.781	0.679
Principal component 3	−29.864	49.378	0.545	7.594	48.627	0.876	38.246	70.134	0.586
Array design 1	−0.043	0.260	0.869	−0.505	0.228	0.027	−0.312	0.279	0.265
Array design 2	−0.268	0.633	0.672	−0.437	0.518	0.399	−0.089	0.713	0.900

Arrays 1 and 2 are two dummy-coded variables included in the model to control for the genotyping arrays. Principal components reflect genetic ancestry. PRS = polygenic risk scores; SE = standard error. Results shown in bold type are significant predictors in the model after multiple testing corrections ($\alpha < 0.0019$).

use comparisons, while thrill-seeking was only significantly associated with belonging to the high trajectory class versus no-low class and with the peer cannabis use comparison between none versus a few (details in Supporting information, Table S6). Cannabis use PRS did not significantly predict either scale, but was significantly associated with conduct disorder diagnosis, as were peer use and all three of the trajectory class comparisons (Supporting information, Table S6). The association between peer use and trajectory membership (high versus no-low class: $P < 0.001$; high versus moderate: $P = 0.037$; moderate versus no-low: $P = < 0.001$) was only somewhat attenuated when including conduct disorder as a covariate. Inclusion of conduct disorder also modestly attenuated the association between PRS and class membership (e.g. high versus no-low class: $OR_{conduct}$ = 1.37, 95% CI = 1.11, 1.69, P = 0.003; versus $OR_{no\text{-}conduct}$: 1.40, 95% CI = 1.13, 1.74, P = 0.002; high versus moderate $OR_{conduct}$ = 1.45, 95% CI = 1.10, 1.90, P = 0.008 versus $OR_{no\text{-}conduct}$ = 1.49, 95% CI = 1.12, 1.97, P = 0.006).

PRS × peer use predicting cannabis use trajectories

PRS × peer use interaction was not significant (Supporting information, Table S7), suggesting independent effects of PRS and peer use on trajectory membership.

© 2018 Society for the Study of Addiction

Table 4 Associations between cannabis initiation polygenic score and perceived peer cannabis use. The PRS that was most strongly associated with cannabis use trajectories is reported ($P_T < 0.1$; see Table 2).

	None versus a few			*None versus most/all*			*A few versus most/all*		
	Beta	*SE*	*P*	*Beta*	*SE*	*P*	*Beta*	*SE*	*P*
Cannabis use PRS	0.021	0.083	0.799	**0.324**	**0.129**	**0.012**	0.184	0.111	0.096
Sex	**0.419**	**0.145**	**0.004**	**0.528**	**0.222**	**0.018**	0.073	0.201	0.717
Age at baseline	**0.099**	**0.045**	**0.026**	**0.315**	**0.072**	**< 0.001**	**0.134**	**0.061**	**0.028**
Age at last assessment	−0.021	0.033	0.529	**−0.124**	**0.052**	**0.018**	−0.055	0.044	0.212
Principal component 1	−380.544	255.575	0.137	**−407.265**	**95.435**	**< 0.001**	116.519	309.088	0.706
Principal component 2	−53.450	135.154	0.692	**−261.639**	**125.320**	**0.037**	−309.740	188.027	0.099
Principal component 3	5.692	52.102	0.913	93.302	63.840	0.144	82.689	67.168	0.218
Array design 1	**0.365**	**0.181**	**0.043**	−0.416	0.269	0.122	**−0.631**	**0.226**	**0.005**
Array design 2	0.402	0.382	0.292	−0.846	0.672	0.208	−0.715	0.582	0.219

Arrays 1 and 2 are two dummy-coded variables included in the model to control for the genotyping arrays. Principal components reflect genetic ancestry. Results shown in bold type are significant predictors in the model ($\alpha < 0.05$).

DISCUSSION

There are three key implications from our study. First, we found a statistically significant association between cannabis PRS and trajectory membership, and the effect size (Δconditional R^2 up to 3.6%) was consistent with other PRS analyses [21]. Thus, genetic propensity to cannabis initiation derived from a large, heterogeneous discovery sample appears to differentiate between classes derived from frequency of cannabis use in an ascertained, longitudinal cohort. Interestingly, life-time cannabis use was not significantly related to PRS. However, maximum frequency of use and DSM-5 CUD were associated with PRS in the larger sample of 1840. It is possible that even though the discovery GWAS was aimed at assessing genetic propensity to life-time use, that polygenic liability is better captured along a developmental spectrum in these data. While, to some extent, the classes differed in severity of use (e.g. CUD), associations with class membership (e.g. high versus no-low) far exceeded cross-sectional associations with CUD, suggesting that class membership in this young and ascertained sample may be a superior index of genetic propensity than cross-sectional indices alone.

Secondly, the 'environmental' risk factor in our study, perceived peer cannabis use, explained up to 11.3% of the variance in trajectory membership. This suggests that, although genetics certainly plays a role in the progression of cannabis use, established environmental influences such as peer use are better predictors of cannabis use than PRS at the moment, and this is also likely to be true for other complex behavioral traits. Uniquely, genetic propensity to cannabis use was also associated with greater perceived peer engagement in cannabis use. Consistent with prior heritability studies, this finding of genetic contributions to perceived peer use might reflect gene–environment correlations [48,49] or causal processes, such as Mendelian randomization [50]. However, both PRS and peer use remained significantly associated with class membership when simultaneously modeled, suggesting some independent effects.

Thirdly, we found no evidence that peer cannabis use is a moderator of polygenic contributions to cannabis use trajectories. Previous studies have found some evidence for interaction effects between peer substance use and genetic liabilities for substance use [48], but few have used genome-wide PRS to do so.

Although results from the discovery GWAS for cannabis use were genetically correlated with risk-taking (SNP-r_g = 0.425, P = 3.4e–42) [19], we found no evidence that our measures of risk-taking were consistently related to the cannabis use PRS. Even though PRS were correlated with conduct disorder, associations between the PRS and trajectory membership persisted even after controlling for conduct disorder. Thus, general deviance does not appear to fully account for these associations.

Our study had several limitations, including a modest target sample size (target n = 1167, discovery sample size n = 162 082; given the current sample size and a significance level of α = 0.05, our study had 80% power [51] to detect an effect size of $R^2 \geq$ 0.0068.). Further replication studies in larger, independent samples are warranted. Also, the current analyses were restricted to individuals of European ancestry, so we cannot confidently extrapolate our conclusions to other populations. Thirdly, COGA is ascertained for genetic liability to addiction, which may have influenced findings. Our 'high' group (16%) is somewhat larger than those noted in two prior general-population longitudinal studies [6,8], but similar to one study that

© 2018 Society for the Study of Addiction

oversampled for tobacco [4] smoking and lower than a study with over-representation of individuals from high crime neighborhoods [7]. Thus, similar classes have been noted, although there is much variability in their class size. Fourthly, while self-report of perceived peer use is commonly studied and does not significantly differ from actual peer use [52], it is possible that it is less objective than reports by peer nominees [53]. Furthermore, as we did not have reports of concurrent peer cannabis use at older ages (and the sample has a diverse age range at final assessment), we cannot speculate whether trajectory membership was associated with subsequent affiliations with cannabis-using peers. Fifthly, we binned frequency of use data into 20-unit intervals and this may have obscured the identification of smaller classes. For instance, our method combined those using one to two times in the past year with those who may have used cannabis 15–20 times. However, sensitivity analyses with 10-unit intervals provided similar results. It is also possible that reported frequency at the upper end of use was imprecise (e.g. using 550 versus 600 times).

It is hoped that with larger discovery efforts of both cannabis use [54] and of cannabis use disorders the predictive quality of PRS, not merely in terms of what they predict, but also when and how they do so, will be elucidated more clearly. However, this study highlights that even as discovery GWAS sample sizes grow and PRS begin to attain a greater level of precision [21,39], it will be of paramount importance to consider not only how genetic liability shapes health and behavior, but also the environmental context within which such behavior unfolds (e.g. [55]).

Declaration of interests

None.

Acknowledgements

This research is supported by the National Institute on Drug Abuse [K02DA32573 (A.A.), DA040411 (E.C.J.), DA040716 (A.P.A.), and K01DA037914 (J.L.M.)], the National Institute on Alcohol Abuse and Alcoholism [K02AA018755 (D.M.D.)], and the National Institute of Mental Health [MH109532 (E.C.J.)]. The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators (PIs) B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes 11 different centers: University of Connecticut (V. Hesselbrock); Indiana University (H. J. Edenberg, J. Nurnberger Jr, T. Foroud); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks); Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia; Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA (L. Almasy), Virginia Commonwealth University (D. Dick), Icahn School of Medicine at Mount Sinai (A. Goate) and Howard University (R. Taylor). Other COGA collaborators include: L. Bauer (University of Connecticut); J. McClintick, L. Wetherill, X. Xuei, Y. Liu, D. Lai, S. O'Connor, M. Plawecki, S. Lourens (Indiana University); G. Chan (University of Iowa; University of Connecticut); J. Meyers, D. Chorlian, C. Kamarajan, A. Pandey, J. Zhang (SUNY Downstate); J.-C. Wang, M. Kapoor, S. Bertelsen (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, S. Saccone (Washington University); J. Salvatore, F. Aliev, B. Cho (Virginia Commonwealth University); and Mark Kos (University of Texas Rio Grande Valley). A. Parsian and H. Chen are the NIAAA Staff Collaborators. We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). This study included summary statistics of a genetic study on cannabis use (Pasman *et al.*, in press, *Nature Neuroscience*). We would like to acknowledge all participating groups of the International Cannabis Consortium, and in particular the members of the working group including Joelle Pasman, Karin Verweij, Nathan Gillespie, Eske Derks and Jacqueline Vink. Pasman *et al.* (2018) included data from the UK Biobank resource under application numbers 9905, 16406 and 25331.

Authors' affiliations

Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA,[1] Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA,[2] Department of Actuarial and Risk Management, Faculty of Business, Karabuk University, Turkey,[3] Department of Psychiatry, SUNY Downstate Medical Center, Brooklyn, NY, USA,[4] Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA,[5] Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA,[6] Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, IA, USA,[7] Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA,[8] Department of Psychiatry, University of California San Diego Medical School, San Diego, CA, USA[9] and Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, NJ, USA[10]

References

1. Arnott D. The impact diverging public opinion on cannabis and tobacco regulation has on constructive engagement between advocates. *Addiction* 2018; **113**: 605–6.

© 2018 Society for the Study of Addiction

2. Hasin D. S., Sarvet A. L., Cerdá M., Keyes K. M., Stohl M., Galea S. *et al*. US adult illicit cannabis use, cannabis use disorder, and medical marijuana laws: 1991–1992 to 2012–2013. *JAMA Psychiatry* 2017; **74**: 579–88.
3. Hasin D. S., Kerridge B. T., Saha T. D., Huang B., Pickering R., Smith S. M. *et al*. Prevalence and correlates of DSM-5 cannabis use disorder, 2012–2013: findings from the National Epidemiologic Survey on Alcohol and Related Conditions–III. *Am J Psychiatry* 2016; **173**: 588–99.
4. Passarotti A. M., Crane N. A., Hedeker D., Mermelstein R. J. Longitudinal trajectories of marijuana use from adolescence to young adulthood. *Addict Behav* 2015; **45**: 301–8.
5. D'Amico E. J., Tucker J. S., Miles J. N. V., Ewing B. A., Shih R. A., Pedersen E. R. Alcohol and marijuana use trajectories in a diverse longitudinal sample of adolescents: examining use patterns from age 11 to 17 years. *Addiction* 2016; **111**: 1825–35.
6. Brook J. S., Zhang C., Brook D. W. Developmental trajectories of marijuana use from adolescence to adulthood: personal predictors. *Arch Pediatr Adolesc Med* 2011; **165**: 55–60.
7. Epstein M., Hill K. G., Nevell A. M., Guttmannova K., Bailey J. A., Abbott R. D. *et al*. Trajectories of marijuana use from adolescence into adulthood: environmental and individual correlates. *Dev Psychol* 2015; **51**: 1650–63.
8. Windle M., Wiesner M. Trajectories of marijuana use from adolescence to young adulthood: predictors and outcomes. *Dev Psychopathol* 2004; **16**: 1007–27.
9. Taylor M., Collin S. M., Munafò M. R., MacLeod J., Hickman M., Heron J. Patterns of cannabis use during adolescence and their association with harmful substance use behaviour: findings from a UK birth cohort. *J Epidemiol Community Health* 2017; **71**: 764–70.
10. Juon H. S., Fothergill K. E., Green K. M., Doherty E. E., Ensminger M. E. Antecedents and consequences of marijuana use trajectories over the life course in an African American population. *Drug Alcohol Depend* 2011; **118**: 216–23.
11. Kandel D. B., Chen K. Types of marijuana users by longitudinal course. *J Stud Alcohol* 2000; **61**: 367–78.
12. Sherva R., Wang Q., Kranzler H., Zhao H., Koesterer R., Herman A. *et al*. Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. *JAMA Psychiatry* 2016; **73**: 472–80.
13. Stringer S., Minică C. C., Verweij K. J., Mbarek H., Bernard M., Derringer J. *et al*. Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. *Transl Psychiatry* 2017; **6**: e769.
14. Demontis D., Rajagopal M. R., Thorgeirsson T. T., Als T. D., Grove J., Pallessen J., *et al*. Genome-wide association study implicates CHRNA2 in cannabis use disorder. *bioRxiv* 2017; https://doi.org/10.1101/237321.
15. Agrawal A., Lynskey M. T., Hinrichs A., Grucza R., Saccone S. F., Krueger R. *et al*. A genome-wide association study of DSM-IV: cannabis dependence. *Addict Biol* 2011; **16**: 514–8.
16. Agrawal A., Lynskey M. T., Bucholz K. K., Kapoor M., Almasy L., Dick D. M. *et al*. DSM-5 cannabis use disorder: a phenotypic and genomic perspective. *Drug Alcohol Depend* 2014; **134**: 362–9.
17. Verweij K. J. H., Vinkhuyzen A. A. E., Benyamin B., Lynskey M. T., Quaye L., Agrawal A., *et al*. The genetic aetiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. *Addict Biol* 2013; **18**: 846–50.
18. Minică C. C., Dolan C. V., Hottenga J. J., Pool R., Genome of the Netherlands Consortium, Fedko I. O., *et al*. Heritability, SNP- and gene-based analyses of cannabis use initiation and age at onset. *Behav Genet* 2015; **45**: 503–13.
19. Pasman J. A., Verweij K. J. H., Gerring Z., Stringer S., Sanchez-Roige S., Treur J. L. *et al*. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. *Nat Neurosci* 2018; **21**: 1161–70.
20. Purcell S. M., Wray N. R., Stone J. L., Visscher P. M., O'Donovan M. C., Sullivan P. F., *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **10**: 8192–8192.
21. Wray N. R., Lee S. H., Mehta D., Vinkhuyzen A. A. E., Dudbridge F., Middeldorp C. M. Research Review: polygenic methods and their application to psychiatric traits. *J Child Psychol Psychiatry Allied Discip* 2014; **55**: 1068–87.
22. Bogdan R., Baranger D. A. A., Agrawal A. Polygenic risk scores in clinical psychology: bridging genomic risk to individual differences. *Annu Rev Clin Psychol* 2018; **14**: 119–57.
23. Adorjan K., Papiol S., Gade K., Malzahn D., Budde M., Aldinger F. *et al*. Polygenic risk scores and substance abuse comorbidity in patients with schizophrenia and bipolar disorders. *Eur Neuropsychopharmacol* 2018; **27**: S409.
24. Kuntsche E., Jordan M. D. Adolescent alcohol and cannabis use in relation to peer and school factors, results of multilevel analyses. *Drug Alcohol Depend* 2006; **84**: 167–74.
25. Whitesell, N. R., Asdigian N. L., Kaufman C. E., Big Crow C., Shangreau C., Keane E. M. *et al*. Trajectories of substance use among young American Indian adolescents: patterns and predictors. *J Youth Adolesc* 2014; **43**: 437–53.
26. Gillespie N. A., Neale M. C., Jacobson K., Kendler K. S. Modeling the genetic and environmental association between peer group deviance and cannabis use in male twins. *Addiction* 2009; **104**: 420–9.
27. Mathys C., Burk W. J., Cillessen A. H. N. Popularity as a moderator of peer selection and socialization of adolescent alcohol, marijuana, and tobacco use. *J Res Adolesc* 2013; **23**: 513–23.
28. Ali M. M., Amialchuk A., Dwyer D. S. The social contagion effect of marijuana use among adolescents. *PLoS ONE* 2011; **6**: e16183.
29. Dishion T. J., Tipsord J. M. Peer contagion in child and adolescent social and emotional development. *Annu Rev Psychol* 2011; **62**: 189–214.
30. Edwards A. C., Maesr H. H., Prescott C. A., Kendler K. S. Multiple mechanisms influencing the relationship between alcohol consumption and peer alcohol use. *Alcohol Clin Exp Res* 2015; **39**: 324–32.
31. Reich, T., Edenberg H. J., Goate A., Williams J. T., Rice J. P., van Eerdewegh P. *et al*. Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* 1998; **81**: 207–15.
32. Bucholz K. K., McCutcheon V. V., Agrawal A., Dick D. M., Hesselbrock V. M., Kramer J. R., *et al*. Comparison of parent, peer, psychiatric, and cannabis use influences across stages of offspring alcohol involvement: evidence from the COGA prospective study. *Alcohol Clin Exp Res* 2017; **41**: 359–68.

© 2018 Society for the Study of Addiction

33. Bucholz, K. K., Cadoret R., Cloninger C. R., Dinwiddie S. H., Hesselbrock V. M., Nurnberger J. I. Jr *et al*. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol* 1994; **55**: 149–58.
34. Hesselbrock M., Easton C., Bucholz K. K., Schuckit M., Hesselbrock V. A validity study of the SSAGA—a comparison with the SCAN. *Addiction* 1999; **94**: 1361–70.
35. American Psychiatric. Association *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. Virginia, PA: American Psychiatric Association Publishing; 2013.
36. So H.-C., Sham P. C. Improving polygenic risk prediction from summary statistics by an empirical Bayes approach. *Sci Rep* 2017; **7**: 41262.
37. Gratten J., Wray N. R., Keller M. C., Visscher P. M. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci* 2014; **17**: 782–90.
38. Wray N. R., Yang J., Hayes B. J., Price A. L., Goddard M. E., Visscher P. M. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet* 2013; **14**: 507–15.
39. Dudbridge F. Power and predictive accuracy of polygenic risk Scores. *PLOS Genet* 2013; **9**: e1003348.
40. Price A. L., Patterson N. J., Plenge R. M., Weinblatt M. E., Shadick N. A., Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–9.
41. Muthén L., Muthén B. *Mplus User's Guide, version 7.4*. Los Angeles: Author; 2012; https://doi.org/10.1111/j.1600-0447.2011.01711.x.
42. Keller M. C. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol Psychiatry* 2014; **75**: 18–24.
43. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Core Team; 2017.
44. Hartz, S. M., Horton A. C., Oehlert M., Carey C. E., Agrawal A., Bogdan R. *et al*. Association between substance use disorder and polygenic liability to schizophrenia. *Biol Psychiatry* 2017; **82**: 709–15.
45. Asparouhov T., Muthén B. Auxiliary variables in mixture modeling: using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes* 2014; **21**: 1–22.
46. Zuckerman M., Kolin E. A., Price L., Zoob I. Development of a sensation-seeking scale. *J Consult Psychol* 1964; **28**: 477–82.
47. Russo, M. F., Stokes G. S., Lahey B. B., Christ M. A. G., McBurnett K., Loeber R. *et al*. A sensation seeking scale for children: further refinement and psychometric development. *J Psychopathol Behav Assess* 1993; **15**: 69–86.
48. Harden K. P., Hill J. E., Turkheimer E., Emery R. E. Gene–environment correlation and interaction in peer effects on adolescent alcohol and tobacco use. *Behav Genet* 2008; **38**: 339–47.
49. Eaves L., Last K., Martin N. G., Jinks J. L. A progressive approach to non-additivity and genotype–environmental covariance in the analysis of human differences. *Br J Math Stat Psychol* 1977; **30**: 1–42.
50. Smith G. D., Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; **32**: 1–22.
51. Erdfelder E., Faul F., Buchner A., Lang A. G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009; **41**: 1149–60.
52. Deutsch A. R., Chernyavskiy P., Steinley D., Slutske W. S. Measuring peer socialization for adolescent substance use: a comparison of perceived and actual friends' substance use effects. *J Stud Alcohol Drugs* 2015; **76**: 267–77.
53. Watson C. G. Do alcoholics give valid self-reports? *J Stud Alcohol* 1984; **45**: 344–8.
54. Pasman J. A., Verweij K. J. H., Gerring Z., Stringer S., Sanchez-Roige S., Treur J. L. *et al*. Genome-wide association analysis of lifetime cannabis use (N=184,765) identifies new risk loci, genetic overlap with mental health, and a causal influence of schizophrenia on cannabis use. *bioRxiv* 2018; https://doi.org/10.1101/234294.
55. Paksarian D., Trabjerg B. B., Merikangas K. R., Mors O., Børglum A. D., Hougaard D. M. *et al*. The role of genetic liability in the association of urbanicity at birth and during upbringing with schizophrenia in Denmark. *Psychol Med* 2018; **48**: 305–14.

[1] PRS = polygenic risk scores; SNP = single nucleotide polymorphism; SE = standard error

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Characteristics of individuals in each subset of data used for analyses.
Table S2 Zero-Inflated Poisson Growth Mixture Models of Cannabis Frequency Category in the Last 12 Months by Age at Assessment (12–30 years) in 1167 European-American Subjects with GWAS and longitudinal data.
Table S3 Parameter estimates for intercepts and slopes for frequency data as well as zero-inflation across 3 latent classes in self-reported European-American subjects.
Table S4 Associations between cannabis use polygenic scores and (a) lifetime cannabis use, (b) frequency of use at baseline, (c) frequency of use at last assessment for individuals, and (d) DSM5 cannabis use disorder diagnosis. The most significant PRS is reported, which was defined with a p-value threshold of $p_T < 0.1$.
Table S5 Mean polygenic risk score across classes derived from the BCH model (covariates include sex, age at first and age at last interview, principal components and recruitment site).
Table S6 Associations between externalizing behaviors and cannabis trajectory membership, peer use, and cannabis use PRS. Trajectory membership and perceived peer use were treated as the dependent variables in these analyses (e.g. peer use regressed on thrill-seeking score), while the PRS was entered as a predictor in the model (e.g. regressing thrill-seeking on PRS).

[1]Arrays 1 and 2 are two dummy-coded variables included in the model to control for the genotyping array types. Principal components reflect genetic ancestry.

© 2018 Society for the Study of Addiction

Table S7 Interaction of polygenic risk scores (PRS) for cannabis initiation and peer cannabis use: results from mixed effect logistic regression models predicting cannabis use trajectory class membership. Models included all PRS*covariate and peer use*covariate interactions.

Figure S1 Three Latent Class Zero-Inflated Poisson Growth Mixture Model of Cannabis Frequency Category in the Last 12 Months by Age at Assessment

Figure S2 Percentage of the variance in cannabis use trajectory class membership explained by cannabis initiation polygenic risk scores defined at varying *p*-value thresholds. The change in conditional R^2 explained by the PRS was calculated for each mixed model (a separate model for each PRS *p*-value threshold). Asterisk indicates the PRS threshold with the most significant association between the PRS and cannabis use trajectory membership.

© 2018 Society for the Study of Addiction